Title	Kolmogorov complexity and the second incompleteness theorem(Mathematical Incompleteness in Arithmetic)
Author(s)	KIKUCHI, MAKOTO
Citation	数理解析研究所講究録 (1995), 912: 33-42
Issue Date	1995-05
URL	http://hdl.handle.net/2433/59565
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
912巻 1995年 33-42

Kolmogorov complexity and the second incompleteness theorem

東北大・理　菊池 誠 (MAKOTO KIKUCHI)

ABSTRACT. It is well known that Kolmogorov complexity has a close relation with Gödel's first incompleteness theorem. In this paper, we give a new formulation of the first incompleteness theorem in terms of Kolmogorov complexity, that is a generalization of Kolmogorov's theorem, and derive a semantic proof of the second incompleteness theroem from it.

Introduction

Kolmogorov complexity is a measure of the quantity of information in finite objects. Roughly speaking, the Kolmogorov complexity of a number n, denoted by $K(n)$, is the size of a program which generates n, and n is called random if $n \leq K(n)$. Kolmogorov showed in 1960's that the set of nonrandom numbers is *simple* in the sense of recursion theory, and this is a version of Gödel's first incompleteness theorem (cf. Odifreddi [Od]). Chaitin also gave another information-theoretic formulation of the first incompleteness theorem in terms of Kolmogorov complexity. Relations between Kolmogorov complexity and the first incompleteness theorem have been discussed in many places (cf. Li and Vitányi [LV]).

Our purpose is to show that Kolmogorov complexity brings the second incompleteness theorem. The common proofs of the first incompleteness theorem by means of Kolmogorov complexity do not yield the second incompleteness theorem in the similar way as the Gödel's argument. Hence, we give a new formulation of the first incompleteness theorem in terms of Kolmogorov complexity by generalizing Kolmogorov's theorem, and derive a semantic proof of the second incompleteness theorem from it.

In spite of their syntactic nature, Gödel's theorems have some semantic proofs. By using models of arithmetic, Kreisel derived new proofs of Gödel's theorems from the arithmetized completeness theorem (cf. Kreisel [Kr] and Smoryński [Sm]), and Paris and Harrington succeeded to give a mathematical (that is, not metamathematical) independent statement, now known as Paris-Harrington principle (see Hájek and Pudlák [HP] and Kaye [Ka]). Recently, Jech [Je] gave a short proof of the second incompleteness theorem by using models of set theory, and Kikuchi [Ki] showed that a formalization of Berry's paradox leads to a model-theoretic proof of the second

incompleteness theorem. Our proof depends on these discussions about models of arithmetic.

In §1, we review basic definitions and theorems in arithmetic and recursion theory, and in §2, define Kolmogorov complexity following Odifreddi [Od]. Then, in §3, we give a proof of the first incompleteness theorem based on Kolmogorov complexity. §4 is devoted to exhibit the arithmetized completeness theorem, which is one of the main tools used in our proof. Finally, in §5, we give a model-theoretic proof of the second incompleteness theorem.

1. Preliminaries

Let $\mathcal{L}_A = \{+, \cdot, 0, 1, <\}$ be the first-order language of arithmetic. Peano arithmetic, denoted by PA, is the theory in $\mathcal{L}_A$ that consists of the basic axioms of arithmetic (i.e., the axioms of discretely ordered semirings with the least positive element 1) and the axiom schema of induction.

We say a quantifier is bounded if it appears in the form $(\forall x < t)\phi$ or $(\exists x < t)\phi$ with a term t that does not contain x. (Here, $(\forall x < t)\phi$ and $(\exists x < t)\phi$ are abbreviations of $\forall x(x < t \to \phi)$ and $\exists x(x < t \wedge \phi)$, respectively.) Then, a formula ϕ in $\mathcal{L}_A$ is called Δ_0 if every quantifier in ϕ is bounded, and called Σ_1 if ϕ is a formula of the form $\exists \bar{x} \psi$ for some Δ_0 formula ψ. It is well known that a relation $R \subseteq \mathbb{N}^k$ is recursively enumerable if and only if there is a Σ_1 formula $\phi(\bar{x})$ such that $\bar{m} \in R$ if and only if $\mathbb{N} \models \phi(\bar{m})$ for all $\bar{m} \in \mathbb{N}^k$.

PA is said to be ω-consistent when the following condition holds: for any formula $\phi(x)$ in $\mathcal{L}_A$, PA $\not\vdash \exists x \neg\phi(x)$ if PA $\vdash \phi(n)$ for all $n \in \mathbb{N}$. Let $\mathrm{Pr}(x)$ be a Σ_1 formula that denotes the relation that *x is the Gödel number of a formula that is derivable from PA*. Then, define Con(PA) and ω-Con(PA) to be sentences in $\mathcal{L}_A$ that mean PA is consistent and ω-consistent respectively.

Theorem 1.1. *For any Σ_1 sentence ϕ in $\mathcal{L}_A$,*
(i) $\mathbb{N} \models \phi \to Pr(\ulcorner \phi \urcorner)$,
(ii) $\mathbb{N} \models Con(PA) \to (Pr(\ulcorner \neg\phi \urcorner) \to \neg\phi)$,
(iii) $\mathbb{N} \models \omega\text{-}Con(PA) \to (Pr(\ulcorner \phi \urcorner) \to \phi)$.

This theorem is provable in PA. That is,

Theorem 1.2. *For any Σ_1 sentence ϕ in $\mathcal{L}_A$,*
(i) $PA \vdash \phi \to Pr(\ulcorner \phi \urcorner)$,
(ii) $PA \vdash Con(PA) \to (Pr(\ulcorner \neg\phi \urcorner) \to \neg\phi)$,
(iii) $PA \vdash \omega\text{-}Con(PA) \to (Pr(\ulcorner \phi \urcorner) \to \phi)$.

See Smoryński [Sm] for the proofs of these theorems.

Let $\{\varphi_e^n(\bar{x})\}_{e \in \mathbb{N}}$ be a canonical enumeration of n-ary recursive functions. (We ommit the superscript n if there is no confusion.) We write $\varphi(\bar{x}) \downarrow$ if $\varphi(\bar{x})$ is defined at $\bar{x}$ and write $\varphi(\bar{x}) \uparrow$ otherwise. Also, we write $\varphi(\bar{x}) \simeq \varphi'(\bar{x})$ if both $\varphi(\bar{x})$ and

$\varphi'(\bar{x})$ are undefined or they are defined and $\varphi(\bar{x}) = \varphi'(\bar{x})$. Note that $\varphi(\bar{x}) \downarrow$ can be expressed by a Σ_1 formula since

$$\varphi(\bar{x}) \downarrow \Leftrightarrow \exists y(y = \varphi(\bar{x}))$$

and the graph of $\varphi(\bar{x})$ can be represented by a Σ_1 formula. We also denote $\varphi_e(\bar{x})$ by $\{e\}(\bar{x})$.

In our later discussion, we use the following theorems of recursion theory (cf. Odifreddi [Od]).

Theorem 1.3. (The S-m-n theorem). *Let m, $n \in \mathbb{N}$ and $\bar{x} = x_1, \ldots, x_m$, $\bar{y} = y_1, \ldots, y_n$. Then, there is a primitive recursive function $S_n^m(z, \bar{x})$ such that*

$$\{S_n^m(e, \bar{x})\}(\bar{y}) \simeq \{e\}(\bar{x}, \bar{y})$$

for all $e \in \mathbb{N}$.

Theorem 1.4. (The recursion theorem). *For any recursive function $f(\bar{x}, y)$, there exists $e \in \mathbb{N}$ such that*

$$\{e\}(\bar{x}) \simeq f(\bar{x}, e).$$

Theorem 1.5. (The selection theorem). *For any recursively enumerable relation $R \subseteq \mathbb{N}^k$, there exists a recursive function $f(\bar{x})$ such that*
(i) $f(\bar{m}) \downarrow$ if $(\bar{m}, n) \in R$ for some $n \in \mathbb{N}$,
(ii) $(\bar{m}, f(\bar{m})) \in R$ if $f(\bar{m}) \downarrow$.

2. Kolmogorov complexity

We define Kolmogorov complexity $K(x)$, a function from $\mathbb{N}$ to $\mathbb{N}$, by

$$K(x) = \mu e(\varphi_e(0) \simeq x).$$

Remark that this function is arithmetically definable, and the relation $K(x) \leq y$ is definable by a Σ_1 formula $\exists z \leq y(\varphi_z(0) \simeq x)$.

Now, we say a number x is *random* if $x \leq K(x)$.

Lemma 2.1. *For any a, there exists $b \leq a + 1$ such that $a + 1 \leq K(b)$*

Proof. Let $a \in N$. Consider the set

$$\{x \in \mathbb{N} : \varphi_e(0) \simeq x \text{ for some } e \leq a\}.$$

Let b be the least number which does not belong to this set. Then, $b \leq a + 1$ and $a + 1 \leq K(b)$. □

Corollary 2.2. *There exist infinitely many random numbers.*

Proof. Let $a \in \mathbb{N}$. By Lemma 2.1, there is $b \in \mathbb{N}$ such that $b \leq a+1$ and $a+1 \leq K(b)$. Then, clearly b is random. Since a is arbitrary, it turns out that there exist infinitely many random numbers. □

The set of nonrandom numbers is recursively enumerable since it is definable by a Σ_1 formula. Kolmogorov showed that this set is not recursive. In fact, he proved the following theorem (cf. Li and Vitányi [LV] and Odifreddi [Od]).

Theorem 2.3. (Kolmogorov). *The set of nonrandom numbers is simple, that is, it is recursively enumerable, and its complement is infinite and does not contain any infinite recursively enumerable subset.*

Clearly, any simple set is not recursive. Since Gödel's first incompleteness theorem is derivable from the existence of a set which is recursively enumerable but not recursive, this theorem is a version of the first incompleteness theorem. We shall prove this theorem in the following section. (See Corollary 3.2.)

Remark that the proof of Lemma 2.1 is formalizable in PA. That is,

Lemma 2.4. $PA \vdash \forall x \exists y \leq x+1(x+1 \leq K(y))$.

Proof. It is enough to show

$$\mathrm{PA} \vdash \forall x \exists y \leq x+1 \forall z \leq x(\neg \varphi_z(0) \simeq y).$$

Use induction on x. □

We use this lemma in the proof of the second incompleteness theorem in §5.

3. The first incompleteness theorem

For any Σ_1 formula $R(x,y)$, let $\Gamma_1(R)$ and $\Gamma_2(R)$ be formulas in $\mathcal{L}_A$ defined by

$$\Gamma_1(R) \Leftrightarrow \forall x \forall y(R(x,y) \to y < K(x)),$$
$$\Gamma_2(R) \Leftrightarrow \forall x \forall y \forall z(R(x,y) \wedge z \leq y \to R(x,z)).$$

Lemma 3.1. *Let $R(x,y)$ be a Σ_1 formula which satisfy*

$$\mathbb{N} \models \Gamma_1(R) \wedge \Gamma_2(R).$$

Then there exits $e \in \mathbb{N}$ such that

$$\mathbb{N} \models \forall x \forall y(R(x,y) \to y < e).$$

Proof. By the selection theorem, there is a recursive function $f(x)$ such that

$$\mathbb{N} \models \exists x R(x,b) \Rightarrow f(b) \downarrow,$$
$$f(b) \downarrow \Rightarrow \mathbb{N} \models R(f(b),b)$$

for all $b \in \mathbb{N}$. Then, by the recursion theorem, there exists $e \in \mathbb{N}$ such that

$$\{e\}(0) \simeq f(e).$$

Claim that $f(e) \uparrow$.

In order to prove this claim, suppose that $f(e) \downarrow$ and $a = f(e)$. Then, $\mathbb{N} \models R(a, e)$ by the condition of f, so $\mathbb{N} \models e < K(a)$ since $\mathbb{N} \models \Gamma_1(R)$. On the other hand, $\{e\}(0) = a$ by the choice of e, so $\mathbb{N} \models K(a) \leq e$, contradiction. Therefore $f(e) \uparrow$.

Now, let a, b be numbers which satisfy $\mathbb{N} \models R(a, b)$. Assume that $e \leq b$. Then, $\mathbb{N} \models R(a, e)$ since $\mathbb{N} \models \Gamma_2(R)$. So $f(e) \downarrow$ by the condition of f, and this contradicts the above claim. Hence we have $\mathbb{N} \models \forall x \forall y(R(x, y) \rightarrow y < e)$. □

As a corollary to this theorem, we can easily deduce Kolmogorov's theorem (Theorem 2.3).

Corollary 3.2. *The set of nonrandom numbers is simple.*

Proof. We have already seen that the set of nonrandom numbers is recursively enumerable and its complement is infinite (Corollary 2.2). Let $P \subseteq \mathbb{N}$ be a recursively enumerable set of random numbers. We show that P is finite. Since every recursively enumerable set can be represented by a Σ_1 formula, there is a Σ_1 formula $R(x, y)$ which satisfy

$$R(x, y) \Leftrightarrow x \in P \wedge y < x.$$

It is clear that $\mathbb{N} \models \Gamma_2(R)$. Since P consists of random numbers, $a \leq K(a)$ for any $a \in P$. Thus $\mathbb{N} \models \Gamma_1(R)$. So, by Lemma 3.1, there exists $e \in \mathbb{N}$ such that

$$\mathbb{N} \models \forall x \forall y(R(x, y) \rightarrow y < e).$$

Furthermore, $\mathbb{N} \models R(a, a - 1)$ if $a \in P$, hence $a \leq e$ for all $a \in P$. This means that P is finite. □

Now, we show our version of the first incompleteness theorem by using Lemma 3.1. First, we remark that a Σ_1 formula $\mathrm{Pr}(\ulcorner y < K(x) \urcorner)$ satisfies the condition of Lemma 3.1.

Lemma 3.3. *(i)* $\mathbb{N} \models \mathit{Con(PA)} \rightarrow \Gamma_1(\mathit{Pr}(\ulcorner y < K(x) \urcorner))$,
(ii) $\mathbb{N} \models \Gamma_2(\mathit{Pr}(\ulcorner y < K(x) \urcorner))$.

Proof. Since $y < K(x)$ is a negation of a Σ_1 formula, (i) is a direct consequence of Theorem 1.1 (ii). It is also easy to show (ii), since $\mathbb{N} \models \forall y \forall z(z \leq y \rightarrow \mathrm{Pr}(\ulcorner z \leq y \urcorner))$ by Theorem 1.1 (i) and $\mathrm{Pr}(\ulcorner x \urcorner)$ satisfies the modus ponens. □

Theorem 3.4. (The first incompleteness theorem). *There exists $e \in \mathbb{N}$ such that*
(i) $\mathbb{N} \models \mathit{Con(PA)} \rightarrow \forall x(\neg \mathit{Pr}(\ulcorner e < K(x) \urcorner))$,
(ii) $\mathbb{N} \models \omega\text{-}\mathit{Con(PA)} \rightarrow \forall x(e < K(x) \rightarrow \neg \mathit{Pr}(\ulcorner K(x) \leq e \urcorner))$.

Proof. (i). By Lemma 3.1 and Lemma 3.3,

$$\mathbb{N} \models \mathrm{Con(PA)} \rightarrow \forall x(\mathrm{Pr}(\ulcorner e < K(x) \urcorner) \rightarrow e < e).$$

(ii) is immediate from Theorem 1.1 (iii). □

4. The arithmetized completeness theorem

Let T be a recursively axiomatizable theory in a language $\mathcal{L}$, C be a set of new constants and $\bar{\mathcal{L}} = \mathcal{L} \cup C$. We say a formula $\phi(x)$ in $\mathcal{L}_A$ defines a model of T in a theory S in $\mathcal{L}_A$ if we can prove within S that the set

$$\{\sigma : \sigma \text{ is a sentence in } \bar{\mathcal{L}} \text{ that satisfy } \phi(\ulcorner \sigma \urcorner)\}$$

forms an elementary diagram of a model of T with a universe from C.

Theorem 4.1. (The arithmetized completeness theorem). *There exists a formula $\mathrm{Tr}_T(\ulcorner x \urcorner)$ in $\mathcal{L}_A$ that defines a model of T in $PA + Con(T)$, where $Con(T)$ is a sentence in $\mathcal{L}_A$ that means T is consistent.*

This theorem is proved by writing down $\mathrm{Tr}_T(\ulcorner x \urcorner)$ actually. The point is the fact that any recursively axiomatizable theory has an arithmetically (but may not recursively) definable complete extension.

Let $\mathfrak{M}$ and $\mathfrak{M}'$ be structures for $\mathcal{L}_A$. We say $\mathfrak{M}'$ is an end-extension of $\mathfrak{M}$ and write $\mathfrak{M} \subseteq_e \mathfrak{M}'$ if $\mathfrak{M} \subseteq \mathfrak{M}'$ and

$$a \in \mathfrak{M} \wedge b \in \mathfrak{M}' \setminus \mathfrak{M} \Rightarrow \mathfrak{M}' \models a < b.$$

Note that

$$\mathfrak{M} \models \phi \Rightarrow \mathfrak{M}' \models \phi$$

if $\mathfrak{M} \subseteq_e \mathfrak{M}'$ and ϕ is a Σ_1 formula.

Let T be a recursively axiomatizable extension of PA and $\mathrm{Pr}_T(\ulcorner x \urcorner)$ be a Σ_1 formula which represents the provability of T. We say a model $\mathfrak{M}'$ of T is a definable end-extension of a model $\mathfrak{M}$ of PA and write $\mathfrak{M} \subseteq_d \mathfrak{M}'$ if $\mathfrak{M} \subseteq_e \mathfrak{M}'$ and they satisfy

$$\mathfrak{M} \models \mathrm{Pr}_T(\ulcorner \phi \urcorner) \Rightarrow \mathfrak{M}' \models \phi$$

and

$$\mathfrak{M} \models \mathrm{Tr}_T(\ulcorner \phi \urcorner) \Leftrightarrow \mathfrak{M}' \models \phi$$

for some formula $\mathrm{Tr}_T(\ulcorner x \urcorner)$ in $\mathcal{L}_A$. From Theorem 4.1, we have the following corollary.

Corollary 4.2. *Let $\mathfrak{M}$ be a model of PA. Then, $\mathfrak{M}$ satisfies $Con(T)$ if and only if $\mathfrak{M}$ has a definable end-extension which is a model of T.*

Proof (sketch). It is clear that $\mathfrak{M}$ has no definable end-extension which is a model of T if $\mathfrak{M} \models \neg\mathrm{Con}(T)$. Conversely, if $\mathfrak{M} \models \mathrm{Con}(T)$, take $\mathrm{Tr}_T(\ulcorner x \urcorner)$ as a formula given by Theorem 4.1 and define an $\mathcal{L}_A$ structure $\mathfrak{M}'$ according to $\mathrm{Tr}_T(\ulcorner x \urcorner)$ and $\mathfrak{M}$. Then $\mathfrak{M}'$ forms a model of T such that $\mathfrak{M} \subseteq_d \mathfrak{M}'$.

See Hájek and Pudlák [HP] and Kaye [Ka] for more information about models of arithmetic, and Smoryński [Sm] and Kikuchi and Tanaka [KT] for the proofs of Theorem 4.1 and Corollary 4.2.

Corollary 4.2 has applications to semantic proofs of the incompleteness theorems (cf. Kreisel [Kr], Smoryński [Sm], and Kikuchi [Ki]). The following is an example of such applications, a proof of the second incompleteness theorem due to Jech [Je].

Theorem 4.3. (The second incompleteness theorem).
If PA is consistent, Con(PA) is not derivable from PA.

Proof (by Jech [Je]). Assume that PA is consistent and Con(PA) holds in any model of PA, and let σ be the Gödel sentence which satisfies

$$\mathrm{PA} \vdash \sigma \leftrightarrow \neg \mathrm{Pr}(\ulcorner \sigma \urcorner).$$

Note that $\neg\sigma$ is a Σ_1 formula. We say a model $\mathfrak{M}$ of PA is positive if $\mathfrak{M} \models \sigma$ and negative otherwise. Since PA is consistent, there is a model $\mathfrak{M}_1$ of PA. If $\mathfrak{M}_1$ is positive, then $\mathfrak{M}_1 \models \mathrm{Con}(PA + \neg\sigma)$, so there is a negative model $\mathfrak{M}_2$ of PA such that $\mathfrak{M}_1 \subseteq_d \mathfrak{M}_2$. Otherwise, let $\mathfrak{M}_2 = \mathfrak{M}_1$. By the assumption that Con(PA) holds for any model of PA, we have a model $\mathfrak{M}_3$ of PA such that $\mathfrak{M}_2 \subseteq_d \mathfrak{M}_3$. Since $\mathfrak{M}_2$ is negative, $\mathfrak{M}_2 \models \mathrm{Pr}(\ulcorner \sigma \urcorner)$, hence $\mathfrak{M}_3$ is positive. But $\mathfrak{M}_3$ must satisfy $\neg\sigma$ since $\neg\sigma$ is a Σ_1 formula and $\mathfrak{M}_2 \subseteq_e \mathfrak{M}_3$, contradiction. □

Remark. (i). In Jech [Je], Jech proved the second incompleteness theorem for set theory. The above proof is a restatement of Jech's proof for arithmetic by means of the arithmetized completeness theorem.
(ii). We can show that $\subseteq_d$ satisfies the transitive law, i.e. $\mathfrak{M}_1 \subseteq_d \mathfrak{M}_2$ and $\mathfrak{M}_2 \subseteq_d \mathfrak{M}_3$ imply $\mathfrak{M}_1 \subseteq_d \mathfrak{M}_3$. Jech's original proof uses this fact with one more step of construction of definable end-extensions instead of using the fact that $\mathfrak{M}_2 \subseteq_e \mathfrak{M}_3$.

5. The second incompleteness theorem

First, we give the formalized versions of Lemma 3.1 and Lemma 3.3.

Lemma 5.1. *Let $R(x, y)$ be a Σ_1 formula. Then there exits $e \in \mathbb{N}$ such that*

$$PA + \Gamma_1(R) \wedge \Gamma_2(R) \vdash \forall x \forall y (R(x, y) \rightarrow y < e).$$

Proof. Let f and e be a recursive function and a number which are given in the proof of Lemma 3.1.

In order to prove the selection theorem in the proof of Lemma 3.1, we define f by

$$f(y) \simeq (\mu w(R'((w)_0, y, (w)_1, \ldots, (w)_n)))_0$$

where R' is the Δ_0 formula such that

$$R(x, y) \Leftrightarrow \exists z_1 \cdots \exists z_n R'(x, y, z_1, \ldots, z_n)$$

and $(w)_i$ is the i-th component of w. So we can prove

$$\mathrm{PA} \vdash \exists x R(x, b) \rightarrow f(b) \downarrow$$
$$\mathrm{PA} \vdash \forall x (x = f(b) \rightarrow R(x, b))$$

for all $b \in \mathbb{N}$.

Also, in order to prove the recursion theorem in the proof of Lemma 3.1, we must apply the S-m-n theorem only for a concrete recursive function f, so we can compute the number e in Lemma 3.1 actually from it. In addition, we can prove $\{e\}(0) \simeq f(e)$ in PA. That is,

$$\mathrm{PA} \vdash \forall x(x = \{e\}(0) \leftrightarrow x = f(e)).$$

Hence, in the same way as in the proof of Lemma 3.1, we can prove

$$\mathrm{PA} + \Gamma_1(R) \vdash \forall x(x = f(e) \to e < K(x))$$
$$\mathrm{PA} \vdash \forall x(x = f(e) \to K(x) \leq e).$$

So we have

$$\mathrm{PA} + \Gamma_1(R) \vdash f(e) \uparrow .$$

Also, since $\forall x \forall y(R(x,y) \wedge e \leq y \to R(x,e))$ is derivable from $\mathrm{PA} + \Gamma_2(R)$, we can prove

$$\mathrm{PA} + \Gamma_2(R) \vdash \forall x \forall y(R(x,y) \wedge e \leq y \to f(e) \downarrow).$$

Hence

$$\mathrm{PA} + \Gamma_1(R) \wedge \Gamma_2(R) \vdash \forall x \forall y(R(x,y) \to y < e). \quad \square$$

Lemma 5.2. *(i)* $PA \vdash Con(PA) \to \Gamma_1(Pr(\ulcorner y < K(x) \urcorner))$,
(ii) $PA \vdash \Gamma_2(Pr(\ulcorner y < K(x) \urcorner))$.

Proof. Use Theorem 1.2 instead of Theorem 1.1. $\square$

From these two lemmas, we have the formalized version of the first incompleteness theorem. Its proof also depends on Theorem 1.2.

Theorem 5.3. (The formalized first incompleteness theorem). *There exists* $e \in \mathbb{N}$ *such that*
(i) $PA \vdash Con(PA) \to \forall x(\neg Pr(\ulcorner e < K(x) \urcorner))$,
(ii) $PA \vdash \omega\text{-}Con(PA) \to \forall x(e < K(x) \to \neg Pr(\ulcorner K(x) \leq e \urcorner))$.

Now, we prove the second incompleteness theorem. We use a mechanism which is parallel to the method in Kikuchi [Ki].

Theorem 5.4. *If PA is consistent, there exists a model of PA which does not satisfy Con(PA).*

Proof. Suppose that PA is consistent and any model of PA satisfies Con(PA). Since PA is consistent, there is a model $\mathfrak{M}_0$ of PA. By Lemma 2.4 and the least number principle in PA, there is $a_0 \leq e + 1$ such that

$$\mathfrak{M}_0 \models e < K(a_0) \wedge \forall x < a_0(K(x) \leq e).$$

Then, by Theorem 5.3 (i),

$$\mathfrak{M}_0 \models \neg \mathrm{Pr}(\ulcorner e < K(a_0) \urcorner).$$

Hence $\mathfrak{M}_0 \models \mathrm{Con}(PA + K(a_0) \leq e)$. So, by Corollary 4.2, there is a definable end-extension $\mathfrak{M}_1$ of $\mathfrak{M}_0$. Again, take the least element $a_1 \leq e+1$ such that $\mathfrak{M}_1 \models e < K(a_1)$. Since $K(x) \leq y$ is a Σ_1 formula, $\mathfrak{M}_1 \models K(a) \leq e$ for any $a < a_0$, and $\mathfrak{M}_1 \models k(a_0) \leq e$ because $\mathfrak{M}_1$ is a model of $PA + K(a_0) \leq e$. Therefore,

$$\mathfrak{M}_1 \models \forall x \leq a_0 (K(x) \leq e),$$

so a_1 is strictly greater than a_0. Repeating this construction $e+2$ times, we have a sequence of models $\mathfrak{M}_0 \subseteq_d \mathfrak{M}_1 \subseteq_d \cdots \subseteq_d \mathfrak{M}_{e+2}$ of PA and a corresponding strictly increasing sequence of numbers $a_0 < a_1 < \cdots < a_{e+2}$. This contradicts the choice of a_i's. □

By the completeness theorem, we have the second incompleteness theorem.

Corollary 5.5. (Gödel's second incompleteness theorem).
If PA is consistent, Con(PA) is not derivable from PA.

Remark. Since our proof of the second incompleteness theorem is not formalizable in the system of primitive recursive arithmetic PRA (cf. Smoryński [Sm], Comments 6.3), it does not directly bring the formalized version of the second incompleteness theorem,

$$\mathrm{PRA} \vdash \mathrm{Con(PA)} \rightarrow \neg \mathrm{Pr}(\ulcorner \mathrm{Con(PA)} \urcorner).$$

However, our proof can be carried out within a subsystem of second-order arithmetic WKL_0, since the completeness theorem is provable in WKL_0 (cf. Simpson [Si]) and Corollary 4.2 is provable in weaker subsystem RCA_0 (cf. Kikuchi and Tanaka [KT]). Thus we can also obtain a new proof of the formalized second incompleteness theorem, by using a theorem of H. Friedman that any Π_2 theorem of WKL_0 is provable in PRA (cf. Simpson [Si]).

Acknowledgement. The author of this paper would like to express his thanks to Prof. K. Tanaka and Mr. T. Yamaguchi for stimulating discussions.

References

[HP] Hájek, P. and Pudlák, P., *Metamathematics of First-Order Arithmetic*, Springer, 1993.
[Je] Jech, T., *On Gödel's second incompleteness theorem*, Proc. Amer. Math. Soc. **121** (1994), 311–313.
[Ka] Kaye, R., *Models of Peano Arithmetic*, Oxford, 1991.
[Ki] Kikuchi, M., *A note on Boolos' proof of the incompleteness theorem*, Math. Logic Quart. **40** (1994), 528–532.
[KT] Kikuchi, M. and Tanaka, K., *On formalization of model-theoretic proofs of Gödel's theorems*, Notre Dame J. Formal Logic (to appear).

[Kr] Kreisel, G., *Notes on arithmetical models for consistent formulae of the predicate calculus*, Fund. Math. **37** (1950), 265–285.
[LV] Li, M. and Vitányi P.M.B., *Kolmogorov complexity and its applications*, Handbook of Theoretical Computer Science (J. van Leeuwen, ed.), Elsevier, 1990, pp. 187–254.
[Od] Odifreddi, P., *Classical Recursion Theorey*, North-Holland, 1989.
[Si] Simpson, S.G., *Subsystems of Second-Order Arithmetic*, forthcoming.
[Sm] Smoryński, C.A., *The incompleteness theorems*, Handbook of Math. Logic (J. Barwise, ed.), North-Holland, 1977, pp. 821–865.

980-77 仙台市青葉区荒巻字青葉 東北大学大学院 理学研究科 数学専攻
E-mail address: makoto@math.tohoku.ac.jp